·论著·

糖尿病足溃疡复发风险预测模型的构建：基于 Logistic 回归、支持向量机和 BP 神经网络模型

张娟，李海芬，李小曼，姚苗，马惠珍，马强*

【摘要】 **背景** 全球范围内糖尿病足溃疡（DFUs）首次复发与再次复发率逐年上升，且早期复发风险高于远期风险。导致 DFUs 复发的风险因素较多，目前缺乏系统的筛选，因此需要探索 DFUs 复发的危险因素，以便早期识别出复发高危人群。**目的** 探讨 Logistic 回归、支持向量机（SVM）和 BP 神经网络（BPNN）模型在糖尿病足溃疡（DFUs）复发风险中的预测价值。**方法** 选取 2020 年 1 月—2022 年 12 月在宁夏医科大学总医院烧伤整形美容科和伤口造口门诊就诊的 DFUs 患者 390 例作为开发模型的研究对象。根据患者出院后 1 年内 DFUs 是否复发分为复发组 116 例（29.7%）和非复发组 274 例（70.3%）。收集两组患者的一般资料包括社会人口学特征、病史评估和临床病例资料并进行比较，采用糖尿病足部自我管理行为量表（DFSBS）评估患者糖尿病足部自我管理行为，采用慢性病风险感知问卷评估患者 DFUs 风险感知水平。采用多因素 Logistic 回归分析探讨 DFUs 患者出院后 1 年内 DFUs 复发的影响因素；将患者按照 70:30 划分为训练集和测试集，运用 Logistic 回归变量筛选策略，分别建立 Logistic 回归、SVM 和 BPNN 模型；绘制各模型预测 DFUs 复发风险的受试者工作特征（ROC）曲线，采用查准率、召回率、正确率、F1 指数及 ROC 曲线下面积（AUC）比较各模型的预测效能。**结果** 两组 DFUs 患者年龄、BMI、发病后就诊时间、住院时间、吸烟史、足溃疡分级、受累足趾截肢史、溃疡位置在脚底、独居、足部存在行走障碍、溃疡因外伤引起、踝肱指数、糖尿病周围神经病变、下肢动脉粥样硬化、多重耐药菌感染、糖化血红蛋白、白蛋白、足部护理行为、风险感知水平比较，差异均有统计学意义（$P<0.05$）。多因素 Logistic 回归分析结果显示，住院时间〔$OR=0.678$，95%CI（0.557，0.826），$P<0.001$〕、溃疡位置在脚底〔$OR=0.078$，95%CI（0.011，0.541），$P=0.010$〕、独居〔$OR=5.689$，95%CI（2.583，10.726），$P<0.001$〕、足部存在行走障碍〔$OR=3.364$，95%CI（2.742，5.638），$P<0.001$〕、糖尿病周围神经病变〔$OR=3.089$，95%CI（1.156，4.585），$P=0.025$〕、下肢动脉粥样硬化〔$OR=4.033$，95%CI（3.688，9.060），$P<0.001$〕、多重耐药菌感染〔$OR=3.241$，95%CI（1.728，7.361），$P<0.001$〕、糖化血红蛋白〔$OR=0.209$，95%CI（0.065，0.669），$P=0.008$〕、白蛋白〔$OR=2.038$，95%CI（1.515，2.741），$P<0.001$〕、足部护理行为〔$OR=1.965$，95%CI（0.874，3.208），$P=0.014$〕和风险感知水平〔$OR=0.261$，95%CI（0.197，0.825），$P=0.002$〕是 DFUs 患者 1 年内 DFUs 复发的影响因素（$P<0.05$）。Logistic 回归、SVM 和 BPNN 模型在测试集中预测 DFUs 复发风险的正确率分别 82.03%、94.87%、87.17%，AUC 分别为 0.843、0.930、0.800。Logistic 回归、SVM 和 BPNN 模型预测 DFUs 复发风险的 ROC 曲线 AUC 比较，差异有统计学意义（$Z=8.826$，$P<0.05$）；SVM 模型预测 DFUs 复发风险的 ROC 曲线 AUC 高于 Logistic 回归和 BPNN 模型（$Z=5.672$，$P=0.014$；$Z=2.767$，$P<0.001$）。**结论** SVM 模型预测 DFUs 患者出院后 1 年内 DFUs 复发风险的正确率、灵敏度、特异度、AUC 等指标均较好，为相对最优的模型，建议进一步推广应用以验证预测模型的效能。

【关键词】 糖尿病；足溃疡；糖尿病足；复发；Logistic 模型；支持向量机模型；BP 神经网络模型；影响因素分析

【中图分类号】 R 587.1 【文献标识码】 A DOI：10.12114/j.issn.1007-9572.2023.0175

【引用本文】 张娟，李海芬，李小曼，等 . 糖尿病足溃疡复发风险预测模型的构建：基于 Logistic 回归、支持向量机和 BP 神经网络模型［J］. 中国全科医学，2023. DOI：10.12114/j.issn.1007-9572.2023.0175.［Epub ahead of print］.［www.chinagp.net］

ZHANG J，LI H F，LI X M，et al. Construction of recurrence risk prediction model for diabetic foot ulcer on the basis of Logistic regression，support vector machine and BP neural network model［J］. Chinese General Practice，2023.［Epub ahead of print］.

基金项目：宁夏回族自治区自然科学基金项目（2022AAC03515）；宁夏回族自治区重点研发计划项目（2020BEG03025）

750001 宁夏回族自治区银川市，宁夏医科大学总医院烧伤整形美容科
*通信作者：马强，副主任医师；E-mail：10901969@qq.com

本文数字出版日期：2023-05-31

Construction of Recurrence Risk Prediction Model for Diabetic Foot Ulcer on the Basis of Logistic Regression, Support Vector Machine and BP Neural Network Model

ZHANG Juan, LI Haifen, LI Xiaoman, YAO Miao, MA Huizhen, MA Qiang[*]

Department of Burn Plastic Surgery, General Hospital of Ningxia Medical University, Yinchuan 750001, China

[*]*Corresponding author: MA Qiang, Associate Chief Physician; E-mail: 10901969@qq.com*

【Abstract】 Background The rates of first and multiple recurrence of diabetic foot ulcers (DFUs) are increasing annually worldwide, and the risk of early recurrence is higher than the distant recurrence. There are numerous risk factors for DFUs recurrence, and there is a lack of systematic screening. Therefore, there is a need to explore the risk factors for DFUs recurrence in order to identify high-risk population of recurrence at an early stage. **Objective** To explore the predictive value of Logistic regression (LR), support vector machine (SVM), BP neural network model (BPNN) in the recurrence risk of DFUs. **Methods** From January 2020 to December 2022, a total of patients with DFUs attending the burn plastic and wound ostomy outpatient department in General Hospital of Ningxia Medical University were selected as the research objects and divided into the recurrence group (n=116, 29.7%) and non-recurrence group (n=274, 70.3%) according to the recurrence of DFUs within 1 year after discharge. General information was collected and compared between the two groups of patients, including sociodemographic characteristics, medical history assessment and clinical case information. The Diabetes Foot Self-care Behavior Scale (DFSBS) was used to assess the self-management behavior of diabetes foot in patients and chronic diseases risk perception questionnaire was used to assess the risk perception level of DFUs among patients. Multivariable Logistic regression analysis was used to explore the influencing factors of DFUs recurrence in patients within 1 year after discharge. The patients were divided into training and test sets according to the ratio of 70 to 30, the LR, SVM and BPNN recurrence risk prediction models were developed based on Logistic regression variable screening strategy. The receiver operating characteristic (ROC) curves of each model to predict the recurrence risk of DFUs were plotted, and the predictive efficacy of each model was compared using the precision rate, recall rate, accuracy rate, F1 index and area under curve (AUC). **Results** There were significant differences in age, BMI, time to visit after onset, length of hospital stay, history of smoking, classification of diabetic foot ulcers, history of involved toe amputation, sole ulcer, living alone, walking impairment, ulcers due to trauma, ankle-brachial index, diabetic peripheral neuropathy, lower limb atherosclerosis, multidrug-resistant bacteria infection, glycated hemoglobin, albumin, foot care behaviour, level of risk perception in both groups of DFUs patients (P<0.05). Multivariable Logistic regression analysis showed that the length of hospital stay [OR=0.678, 95%CI (0.557, 0.826), P<0.001], sole ulcer [OR=0.078, 95%CI (0.011, 0.541), P=0.010], living alone [OR=5.689, 95%CI (2.583, 10.726), P<0.001], walking impairment [OR=3.364, 95%CI (2.742, 5.638), P<0.001], diabetic peripheral neuropathy [OR=3.089, 95%CI (1.156, 4.585), P=0.025], lower limb atherosclerosis [OR=4.033, 95%CI (3.688, 9.060), P<0.001], multidrug-resistant bacteria infection [OR=3.241, 95%CI (1.728, 7.361), P<0.001], glycated haemoglobin [OR=0.209, 95%CI (0.065, 0.669), P=0.008], albumin [OR=2.038, 95%CI (1.515, 2.741), P<0.001], foot care behaviour [OR=1.965, 95%CI (0.874, 3.208), P=0.014] and level of risk perception [OR=0.261, 95%CI (0.197, 0.825), P=0.002] were influencing factors of the recurrence of DFUs within 1 year (P<0.05). The accuracy rates of LR, SVM and BPNN models to predict the recurrence risk of DFUs in the test sets were 82.03%, 94.87% and 87.17%, with AUCs of 0.843, 0.930 and 0.800, respectively. There were significant differences in AUC of ROC curves of LR, SVM and BPNN recurrence risk prediction models of DFUs (Z=8.826, P<0.05); the AUC of ROC curve of SVM recurrence risk prediction model was higher than the LR and BPNN models (Z=5.672, P=0.014; Z=2.767, P<0.001). **Conclusion** SVM model can predict the recurrence risk of DFUs patients within 1 year after discharge with good accuracy rate, sensitivity, specificity, AUC and other indicators, which is the relative optimal model. It is recommended to further promote and apply the prediction model to verify its effectiveness.

【Key words】 Diabetes mellitus; Foot ulcer; Diabetic foot; Recurrence; Logistic models; Support vector machine; Back propagation neural network; Root cause analysis

根据国际糖尿病协会数据，2019年全球糖尿病患病率为9.3%，到2045年将升至10.9%[1]。与此同时，全球慢性创面的患病率也在增加，以糖尿病足溃疡（diabetic foot ulcers，DFUs）和下肢静脉溃疡最为常见[2]。DFUs是糖尿病最严重的并发症之一，可导致高死亡率和致残率[3]。DFUs患者的终生风险达到15%~25%[4]，DFUs患者截肢的年发生率为5.1%[1]。有研究报道即使创面愈合后DFUs在一年内复发率高达40%，3年复发率为50%~60%，5年内复发率为65%[4-7]。显然，DFUs预后不稳定，易复发且早期复发风险较高。

吕静等[8]采用 Logistic 回归建立 DFUs 复发风险预测模型，1 年内 DFUs 复发率为 26.9%，吸烟、愈伤组织、足部肤色异常、胼胝质、糖尿病周围神经病变和冠心病是 DFUs 复发的风险因素。CRAWFORD 等[9]的研究显示 DFUs 史、无法感受到 10 g 单丝和没有任何足踏板脉搏是 DFUs 复发的独立危险因素。最近，AAN DE STEGGE 等[10]建立了 DFUs 复发风险预测模型，模型 1 主要针对复发性足底溃疡，模型 2 主要针对无法识别的重复性压力引起的溃疡。以往研究中 DFUs 复发风险预测模型的建立方法单一，未对比不同算法的预测效能，在此情况下本研究运用机器学习（machine learning，ML）中的支持向量机（support vector machine，SVM）和 BP 神经网络（back propagation neural network，BPNN）开发出 DFUs 复发风险预测模型，并对比其预测效能，以便临床医护人员能够及时识别出有复发性足溃疡风险的患者，指导临床医生和患者进行预防性治疗的联合决策，最终达到降低 DFUs 复发风险的目的。

1 对象与方法

1.1 研究对象 选取 2020 年 1 月—2022 年 12 月在宁夏医科大学总医院烧伤整形美容科、内分泌科和伤口造口门诊就诊的 DFUs 患者作为开发模型的研究对象。纳入标准：（1）符合《中国老年 2 型糖尿病防治临床指南（2022 年版）》诊断标准[11]的 2 型糖尿病患者；（2）因初发 DFUs 而就诊；（3）自愿参与，且签署知情同意书。排除标准：（1）糖尿病合并妊娠或特殊类型糖尿病；（2）合并严重恶性肿瘤等危重疾病者；（3）有沟通障碍者；（4）有精神疾患者；（5）中途失访者。本研究获批宁夏医科大学总医院伦理审查（编号：KYLL-2021-677），纳入研究对象均知情同意且签署知情同意书。

1.2 样本量计算 本研究共纳入 36 个候选预测变量，根据多变量预测模型透明报告对样本量的要求[14]：阳性结局事件数至少是候选预测变量个数的10倍及以上，即本模型开发至少需要 360 个 DFUs 结局事件。经数据清洗，本研究共纳入 390 例 DFUs 患者为研究对象。本研究结局变量是 DFUs 愈合出院后 1 年内是否复发的情况。DFUs 复发诊断标准参照国际糖尿病足溃疡工作组（IWGDF）指南[12]和《2020 IWGDF 指南：糖尿病足的诊断》标准[13]：既往有过 DFUs 史的患者出现新的足溃疡，不论这次溃疡发生的部位与上次是否相同。随访 1 年，以初诊 DFUs 住院诊治且溃疡愈合后的时间为起点，若随访期内同一患者 DFUs 多次复发，则仅记 1 次，患者复发信息以第一次复发时收集的信息为参考。根据患者出院后 1 年内 DFUs 是否复发分为复发组 116 例（29.7%）和非复发组 274 例（70.3%）。

1.3 观察指标

1.3.1 一般资料 根据研究目的及内容，由研究小组通过文献回顾、预调查和专家建议等方式自行设计一般资料收集表，内容包括社会人口学特征、病史评估、临床病例资料。（1）社会人口学特征包括性别、年龄、BMI、发生 DFUs 后就诊时间、住院时间等；（2）病史评估包括糖尿病病程、吸烟史（每天至少抽 1 支烟，持续或累积半年以上且仍在吸烟[14]）、饮酒史（每周至少饮酒 1 次，持续或累计饮酒半年以上且仍在饮酒[14]）、DFUs 史、高血压病史、冠心病病史、糖尿病周围神经病变、糖尿病视网膜病变、糖尿病肾病史；（3）临床病例资料包括足溃疡分级、足部皮肤颜色、足趾受累、溃疡位置在脚底、足部胼胝、受累足趾截肢史、下肢动脉粥样硬化、独居、足部存在行走障碍、溃疡因外伤引起、踝肱指数、骨髓炎、多重耐药菌感染、糖化血红蛋白、总胆红素、白蛋白、肌酐、胆固醇、血红蛋白、白细胞计数。

1.3.2 糖尿病足部自我管理行为量表（Diabetes Foot Self-care Behavior Scale，DFSBS） 采用 DFSBS 评估糖尿病足部自我管理行为，DFSBS 由靳燕芬等[15]研制，共 7 个条目，其中“我会自己（或请家人帮忙）检查脚底”“我会自己（或请家人帮忙）检查脚趾缝”“洗脚时我会自己（或请家人帮忙）清洗脚趾缝”“洗脚时我会自己（或请家人帮忙）擦干脚趾缝”4 个条目调查患者过去 7d 实施足部自我管理行为的天数，设为 0、1~2、3~4、5~6 和 7 d，共 5 个等级，分别评为 1~5 分，其余 3 个条目调查患者足部管理行为的执行频率，按照“从不”“很少”“有时”“经常”“总是”，分别评为 1~5 分，量表总分为 7~35 分，得分越高表示足部自我管理行为越好，量表 Cronbach's α 系数为 0.835，重测信度为 0.916。

1.3.3 慢性病风险感知问卷 采用慢性病风险感知问卷评估患者 DFUs 风险感知，该问卷由方蕾等[16]编制，用于评估慢性病患者的风险感知，包括经济风险、身体诊疗风险及社会心理风险，共 3 个维度 12 个条目。采用 Likert-5 级评分法评分，从“一定不会发生的”“不太可能发生的”“可能会发生的”“很可能发生的”“一定会发生的”分别计 1~5 分，总分为 12~60 分，得分越高表示研究对象的医疗风险感知水平越高。该问卷的 Cronbach's α 系数为 0.833，各维度 Cronbach's α 系数为经济风险 0.716、身体诊疗风险 0.769，社会心理风险 0.781，具有良好的信效度。

1.4 质量控制 运用医院信息系统收集临床资料，通过电话随访、门诊随访与院外随访管理系统相结合的方式，记录患者 DFUs 的复发情况，随访时间为 1 年，由临床医生、糖尿病专科护士或伤口专科护士判断患者的复发情况及症状的严重程度。收集资料前统一在烧伤整形美容科、伤口造口门诊和内分泌科选派 3 名糖尿病专

科护士和 3 名伤口专科护士共 6 名作为资料收集员，由研究者对其进行同质化培训和指导，包括随访资料收集的内容、方法和注意事项等。选取 5 例 DFUs 患者进行预调查以完善资料收集表。采用问卷星编制《糖尿病足溃疡问卷调查表》，运用金山文档编制《糖尿病足溃疡在线资料收集表》，通过以上两种途径完成资料收集。由培训合格的资料收集员协助患者作答问卷。电话随访由烧伤整形美容科护士进行，门诊随访在伤口造口门诊完成，院外管理系统随访由内分泌科教育护士完成。

1.5 模型构建

1.5.1 数据初步筛选和预处理 审核数据的变量特征，并按既定规则进行初步筛选：（1）删除每列数据缺失占比 >90% 的变量；（2）删除每列单个类别比例 >90% 的变量；（3）删除每列变异系数（variable coefficient，CV）<0.05 的变量；（4）删除不合逻辑的异常值，连续性变量的缺失值用中位数填补，分类变量的缺失值用众数填补。

1.5.2 特征及算法的选择 本研究的特征选择步骤在训练数据集中，以 DFUs 1 年内是否复发为因变量，其他变量为自变量进行初始逐步 Logistic 回归分析，利用 *OR* 值及其 95%*CI* 等指标增加模型的可解释性。将单因素分析中 *P*<0.05 的预测因子纳入多因素分析，最后使用逐步回归确定进入模型的预测因子。BPNN 对数据资料分布无严格限制，可识别变量间复杂的非线性关系，具有很强的自适应能力，且有高度的容错性[17]。SVM 是一种线性和非线性分类方法，具有预测可靠性高、稳定性强和泛化能力强等优势[18]。

1.5.3 模型的建立及评价 纳入 Logistic 回归中有意义的 11 个独立预测因素，构建 Logistic、BPNN 及 SVM 风险预测模型，各模型训练集和测试集按 70 ∶ 30 比例，292 例样本用于训练，98 例样本用于预测。以受试者工作特征曲线（ROC）曲线下面积（AUC）、查准率、正确率、召回率和 F1 值作为模型评估的指标，在各指标结果不一致时，以 AUC 作为主要参考，其中 AUC 为 0.50~0.70 提示预测效果较差，AUC 为 0.70~0.90 提示预测效果中等，AUC>0.9 提示预测效果非常好。

1.6 统计学方法 运用 Excel 导出并录入数据，采用 SPSS 25.0 及 MATLAB R2020b 分析数据。正态分布的定量资料以（$\bar{x} \pm s$）表示，两组间比较采用两独立样本 *t* 检验；非正态分布的定量资料以 M（P_{25}，P_{75}）表示，两组间比较采用独立样本 Mann-Whitney *U* 检验；定性资料以频数及百分比表示，两组间比较采用 χ^2 检验，等级资料比较采用 Wilcoxon W 检验。采用 SPSS 25.0 构建 Logistic 回归模型预测 DFUs 复发的影响因素，采用 MATLAB R2020b 软件构建 BPNN 和 SVM 模型，模型的拟合优度采用 Hosmer - Lemeshow 检验，绘制各模型预测 DFUs 复发风险的 ROC 曲线，采用查准率、正确率、召回率、F1 指数及 AUC 评价各模型的预测效能。运用 DeLong 检验比较各模型的性能及 AUC 间的差异。以 *P*<0.05 为差异有统计学意义。

2 结果

2.1 一般资料比较 两组 DFUs 患者性别、糖尿病病程、饮酒史、足部皮肤颜色、足部胼胝、足趾受累、高血压病史、冠心病病史、糖尿病视网膜病变、糖尿病肾病史、骨髓炎、总胆红素、肌酐、胆固醇、血红蛋白、白细胞计数比较，差异均无统计学意义（*P*>0.05）；两组 DFUs 患者年龄、BMI、发病后就诊时间、住院时间、吸烟史、足溃疡分级、受累足趾截肢史、溃疡位置在脚底、独居、足部存在行走障碍、溃疡因外伤引起、踝肱指数、糖尿病周围神经病变、下肢动脉粥样硬化、多重耐药菌感染、糖化血红蛋白、白蛋白、足部护理行为、风险感知水平比较，差异均有统计学意义（*P*<0.05），见表 1。

2.2 DFUs 患者出院后 1 年内 DFUs 复发影响因素的多因素 Logistic 回归分析 以 DFUs 患者出院后 1 年内 DFUs 是否复发为因变量，以表 1 中差异有统计学意义的因素：年龄、BMI、发病后就诊时间、住院时间、吸烟史、足溃疡分级、受累足趾截肢史、溃疡位置在脚底、独居、足部存在行走障碍、溃疡因外伤引起、踝肱指数、糖尿病周围神经病变、下肢动脉粥样硬化、多重耐药菌感染、糖化血红蛋白、白蛋白、足部护理行为、风险感知水平为自变量进行多因素 Logistic 回归分析（各变量赋值情况见表 2），结果显示，住院时间、溃疡位置在脚底、独居、足部存在行走障碍、糖尿病周围神经病变、下肢动脉粥样硬化、多重耐药菌感染、糖化血红蛋白、白蛋白、足部护理行为和风险感知水平是 DFUs 复发的影响因素（*P*<0.05），见表 3。

2.3 三种模型的运行结果

2.3.1 Logistic 回归模型运行结果 将多因素 Logistic 回归分析有统计学意义的 11 个影响因素作为预测模型的输入变量。Logistic 回归在训练集上的拟合值与真实值相比，其预测 DFUs 患者 DFUs 复发的查准率、正确率和 AUC 分别为 83.26%、81.41% 和 0.825。

2.3.2 SVM 模型运行结果 以 DFUs 是否复发为因变量（赋值：否 =0，是 =1），多因素 Logistic 回归分析筛选出的变量为自变量（赋值同表 2），在训练集上建立 SVM，利用 tune.svm（）函数，通过网格搜索方法，尝试各种参数组合以寻找最优参数，C 与 γ 初始取值范围皆为（0.01，0.1，1，10，100）。10 折交叉验证错误率最低时的最优参数为 C=100、γ=0.01，此时训练集 SVM 预测 DFUs 患者复发的查准率、正确率、AUC 分别为 97.20%、96.33% 和 0.940，见表 4。

2.3.3 BPNN 模型运行结果 以 DFUs 是否复发为因变

表 1 两组 DFUs 患者一般资料比较

Table 1 Comparison of general information between the two groups of patients with DFUs

组别	例数	性别〔例（%）〕		年龄（岁）	BMI（kg/m²）	糖尿病病程（年）	发病后就诊时间（d）	住院时间（d）	吸烟史〔例（%）〕	饮酒史〔例（%）〕
		男	女							
复发组	116	55（47.4）	61（52.6）	58.8±12.9	22.3±1.2	13.4±1.9	7.2±1.5	24.2±3.6	73（62.9）	95（81.9）
未复发组	274	152（55.5）	122（44.5）	57.7±9.2	23.2±1.4	11.4±2.4	8.4±2.4	12.6±5.3	176（59.9）	200（73.0）
检验统计量值		2.126		4.026[a]	−2.298[a]	2.793[a]	−4.870[a]	21.728[a]	2.060	3.507
P 值		0.151		0.004	0.022	0.059	0.002	0.022	0.047	0.061

组别	足溃疡分级〔例（%）〕					受累足趾截肢史〔例（%）〕		溃疡位置在脚底〔例（%）〕	足部皮肤颜色〔例（%）〕	足部胼胝〔例（%）〕
	Ⅰ级	Ⅱ级	Ⅲ级	Ⅳ级	Ⅴ级	差	良			
复发组	4（3.4）	17（14.7）	43（37.0）	35（30.2）	17（14.7）	57（49.1）	61（52.6）	69（59.5）	47（40.5）	56（48.3）
未复发组	62（22.6）	88（32.1）	74（27.0）	43（15.7）	7（2.6）	65（23.7）	106（38.7）	132（48.2）	142（51.8）	131（47.8）
检验统计量值	57.628[b]					24.489		6.431	3.176	2.562
P 值	<0.001					<0.001		0.011	0.053	0.374

组别	足趾受累〔例（%）〕	独居〔例（%）〕	足部存在行走障碍〔例（%）〕	溃疡因外伤引起〔例（%）〕	踝肱指数〔例（%）〕		高血压病史〔例（%）〕	冠心病病史〔例（%）〕	糖尿病周围神经病变〔例（%）〕
					差	良			
复发组	56（48.3）	81（69.8）	103（88.8）	90（77.6）	70（60.3）	46（39.6）	67（57.8）	59（50.9）	86（74.1）
未复发组	132（48.2）	161（58.8）	171（62.4）	154（56.2）	90（32.8）	184（67.2）	130（47.4）	146（53.3）	120（43.8）
检验统计量值	3.819	4.240	27.149	5.571	25.470		3.468	0.192	30.108
P 值	0.072	0.039	<0.001	0.026	<0.001		0.063	0.661	<0.001

组别	糖尿病视网膜病变〔例（%）〕	糖尿病肾病史〔例（%）〕	下肢动脉粥样硬化〔例（%）〕	骨髓炎〔例（%）〕	多重耐药菌感染〔例（%）〕	糖化血红蛋白（%）	总胆红素〔M（P_{25}，P_{75}），μmol/L〕	白蛋白（g/L）
复发组	69（59.5）	40（34.5）	55（47.4）	85（73.3）	86（74.1）	9.51±0.72	9.73（7.70，11.30）	34.47±3.03
未复发组	122（44.5）	79（28.8）	99（36.1）	124（45.2）	72（26.3）	7.62±0.75	9.05（8.30，11.52）	41.34±2.42
检验统计量值	5.564	1.227	4.342	2.403	77.954	23.079[a]	−0.753[c]	−23.825[a]
P 值	0.273	0.268	0.037	0.301	<0.001	<0.001	0.438	<0.001

组别	肌酐〔M（P_{25}，P_{75}），umol/L〕	胆固醇（mmol/L）	血红蛋白〔M（P_{25}，P_{75}），g/L〕	白细胞计数〔M（P_{25}，P_{75}），$\times 10^9$/L〕	足部护理行为〔M（P_{25}，P_{75}），分〕	风险感知水平〔M（P_{25}，P_{75}），分〕
复发组	99.52（98.37，103.77）	6.42±2.06	99.38（88.75，116.73）	10.03（8.37，12.64）	46.93（32.74，52.68）	31.34（28.32，42.37）
未复发组	98.85（97.95，105.72）	6.80±2.14	99.05（91.26，115.73）	10.15（7.53，13.80）	63.83（49.21，82.34）	47.30（39.74，65.52）
检验统计量值	−0.726[c]	0.311[a]	−1.772[c]	−0.426[c]	3.642[c]	−2.146[c]
P 值	0.424	0.568	0.069	0.670	0.006	0.032

注：[a]为 t 值，[b]为 T 值，[c]为 Z 值，余检验统计量值为 χ^2 值。

量（赋值：否 =0，是 =1），以多因素 Logistic 回归分析中有统计学意义的变量为自变量（赋值同表 2），计量资料采用标准化公式，将输入层协变量全部归一化处理为（0，1）（区间的变量）输入 BPNN 模型。对测试集预测的正确率为 87.17%，训练集预测的正确率为 88.64%，见表 4。

2.4 三种模型的预测价值比较 三种模型预测患者 DFUs 复发的 ROC 曲线均具有较高的正确率，在训练集和测试集中，三种模型预测 DFUs 复发风险的 ROC 曲线 AUC 比较结果为：SVM>Logistic 回归 >BPNN。由 ROC 曲线可知，SVM 在训练集和测试集中预测患者 DFUs 复发的 AUC 均最高，测试集中的 AUC 为 0.930〔95%*CI*（0.916，0.955）〕，且正确率最高为 94.87%，灵敏度为 0.99，特异度为 0.87，SVM 模型整体表现较另外两种预测模型预测效能更优。经 DeLong 检验，Logistic 回归、SVM 和 BPNN 模型预测 DFUs 复发风险的 ROC 曲线 AUC 比较，差异有统计学意义（Z=8.826，P<0.05）；SVM 模型预测 DFUs 复发风险的 ROC 曲线 AUC 高于 Logistic 回归和 BPNN 模型，差异有统计学意义（Z=5.672，P=0.014；Z=2.767，P<0.001）；Logistic 回归模型预测 DFUs 复发风险的 ROC 曲线 AUC 高于 BPNN 模型，差异有统计学意义（Z=3.274，P=0.028），见图 1、表 4。

3 讨论

DFUs 复发是糖尿病患者不良预后的重要标志[19]。本研究关注引起 DFUs 复发的主客观变量，自行研制 DFUs 复发高危因素在线筛查表，短期随访（出院后 1 年）DFUs 患者的复发状况，运用机器学习算法探讨 DFUs 复发的风险因素，构建出复发风险预测模型以便实现

表 2 DFUs 复发影响因素的多因素 Logistic 回归分析赋值表

Table 2 Assignment table for multivariable Logistic regression analysis of influencing factors of DFUs recurrence

变量	赋值
DFUs 复发	是 =1，否 =0
年龄	实测值
BMI	实测值
发病后就诊时间	实测值
住院时间	实测值
吸烟史	有 =1，无 =0
足溃疡分级	Ⅰ级 =1，Ⅱ级 =2，Ⅲ级 =3，Ⅳ级 =4，Ⅴ级 =5
受累足趾截肢史	有 =1，无 =0
溃疡位置在脚底	是 =1，否 =0
独居	是 =1，否 =0
足部存在行走障碍	是 =1，否 =0
溃疡因外伤引起	是 =1，否 =0
踝肱指数	良 =0，差 =1
糖尿病周围神经病变	有 =1，无 =0
下肢动脉粥样硬化	有 =1，无 =0
多重耐药菌感染	有 =1，无 =0
糖化血红蛋白	实测值
白蛋白	实测值
足部护理行为	实测值
风险感知水平	实测值

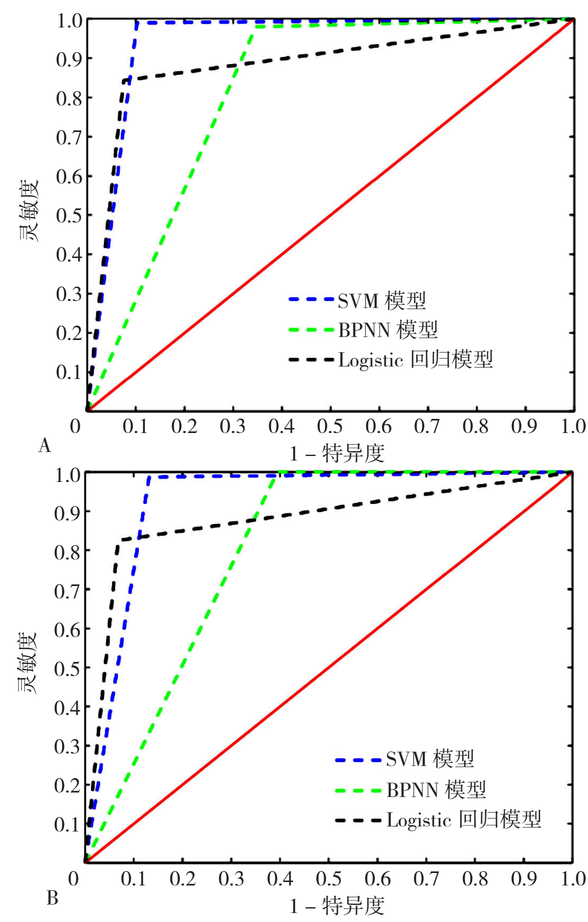
表 3 DFUs 复发影响因素的多因素 Logistic 回归分析

Table 3 Multivariable Logistic regression analysis of influencing factors of DFUs recurrence

自变量	*β*	*SE*	Wald χ^2 值	*P* 值	*OR* 值	95%*CI*
常量	-8.932	1.253	48.637	<0.001	1.005	—
年龄	0.088	0.072	1.507	0.220	0.916	（0.795，1.054）
BMI	0.287	0.376	0.584	0.445	1.333	（0.638，2.783）
发病后就诊时间	0.086	0.233	0.138	0.510	1.090	（0.691，1.720）
住院时间	-0.388	0.101	14.901	<0.001	0.678	（0.557，0.826）
吸烟史	4.147	3.647	1.293	0.256	1.532	（1.250，1.558）
足溃疡分级	—	—	4.104	0.392	—	—
受累足趾截肢史	1.434	0.918	2.437	0.119	0.838	（0.802，1.027）
溃疡位置在脚底	2.550	0.989	6.666	0.010	1.478	（1.239，1.645）
独居	2.535	1.049	5.833	<0.001	5.689	（2.583，10.726）
足部存在行走障碍	1.599	0.249	4.640	<0.001	3.364	（2.742，5.638）
溃疡因外伤引起	0.961	0.700	1.844	0.170	0.382	（0.097，1.509）
踝肱指数	0.578	1.342	0.186	0.667	0.782	（0.715，0.974）
糖尿病周围神经病变	0.985	0.490	6.107	0.025	3.089	（1.156，4.585）
下肢动脉粥样硬化	3.527	1.134	9.678	<0.001	4.033	（3.688，9.060）
多重耐药菌感染	1.799	0.916	4.584	<0.001	3.241	（1.728，7.361）
糖化血红蛋白	1.568	0.595	6.938	0.008	0.209	（0.065，0.669）
白蛋白	-0.712	0.151	22.182	<0.001	2.038	（1.515，2.741）
足部护理行为	-2.324	0.437	52.396	0.014	1.965	（0.874，3.208）
风险感知水平	-1.683	0.193	34.071	0.002	0.261	（0.197，0.825）

注：—表示无此项数据。

快速筛查，旨在为糖尿病高危足 DFUs 复发提供全面的精准干预。本研究 DFUs 患者出院后 1 年内的复发率为 29.74%，高于吕静等[8] DFUs 1 年复发率为 26.9% 的研究结果，低于 WANG 等[20]和 JIANG 等[21]报道的治愈



注：A 图为训练集 ROC 曲线，B 图为测试集 ROC 曲线；SVM=支持向量机，BPNN=BP 神经网络。

图 1 Logistic 回归、SVM 和 BPNN 模型训练集和测试集预测 DFUs 复发风险的 ROC 曲线

Figure 1 ROC curves of Logistic Regression，SVM and BPNN recurrence risk prediction models for DFUs in the the training and test sets

表 4 Logistic 回归、SVM 和 BPNN 模型对 DFUs 复发风险的预测性能比较

Table 4 Comparison of predictive performance of Logistic Regression，SVM and BPNN recurrence risk prediction models for DFUs

评价指标	训练集			测试集		
	Logistic 回归模型	SVM 模型	BPNN 模型	Logistic 回归模型	SVM 模型	BPNN 模型
查准率（%）	83.26	97.20	92.70	84.42	97.06	96.00
召回率（%）	80.95	89.74	65.40	79.16	86.84	60.50
正确率（%）	81.41	96.33	88.64	82.03	94.87	87.17
F1 指数	0.82	0.97	0.92	0.81	0.96	0.91
AUC	0.825	0.940	0.820	0.843	0.930	0.800
灵敏度	0.97	0.99	0.98	0.97	0.99	1.00
特异度	0.81	0.90	0.75	0.78	0.87	0.71

后 DFUs 1 年内再发病率为 31.6%，且低于 HICKS 等[22]的研究结果。目前各区域性医疗机构有关 DFUs 复发率的报道有所不同[23-24]，究其原因与医疗卫生技术、居民生活经济水平、区域性糖尿病防治教育普及程度、糖尿病高危足筛查及精准诊疗等因素有关。提前对高危足进行筛查和防治对预防 DFUs 复发尤为重要[25]。通过数据挖掘技术形成的预测模型有助于在院外尽早识别 DFUs 复发的危险因素，可及早发现 DFUs 好发的患病高危人群，对高危人群实施干预措施，对于有严重复杂并发症的 DFUs，可启动快速转诊路径，送往专业的足病治疗中心，对提高人群健康水平意义重大。

本研究从简单个人水平和复杂临床水平分别筛选出与 DFUs 复发相关的影响因素，运用机器学习算法建立最优预测模型，最后进行模型评价。多因素 Logistic 回归分析结果显示，住院时间、溃疡位置在脚底、独居、足部存在行走障碍、糖尿病周围神经病变、下肢动脉粥样硬化、多重耐药菌感染、糖化血红蛋白、白蛋白、足部护理行为和风险感知水平是 DFUs 患者 1 年内 DFUs 复发的影响因素，与大多数研究结果相符[7，8，11，13，24]，为 DFUs 复发的早期筛查与防治提供了可行干预方案。

目前，许多关于 DFUs 复发风险的研究常使用传统或常见的机器学习算法建立模型，很少采用多种机器学习算法进行多样化建模来比较模型预测性能。本研究预测模型指标筛选共纳入 36 个变量，将多因素 Logistic 回归分析中有统计学意义的 11 个因素纳入模型构建，分别建立 Logistic 回归、SVM 与 BPNN 风险预测模型，并且在测试集中评估了三种模型的预测效能，研究结果显示，三种预测模型 AUC 均 >0.70，正确率均 >80%，表示具有较高的准确性和良好的预测能力[26-27]。训练集与测试集中 SVM 模型的各项评价指标均高于 Logistic 回归和 BPNN 模型，BPNN 相比 Logistic 模型的召回率较低，但查准率更准确。仅用单个指标无法对算法的效果进行总体评价，为调和查准率与召回率的差异，更加准确地评价模型的综合性能[28]，本研究使用 F1 指数对各模型预测效果进行综合评价，其结果为 BPNN>Logistic 回归，可能与样本量和变量数影响预测模型的效能有关系，本研究中测试集的预测数据量有限，限制了 BPNN 算法的优势。综合上述指标，基于 SVM 模型构建的预测模型具有良好诊断效能且更为稳定，在预测 DFUs 复发方面的正确率、灵敏度和 AUC 等性能指标上优于其他两种模型，提示通过结合临床病例资料和患者报告结局资料所建立的预测模型可作为有效预测 DFUs 复发的辅助决策工具，具有良好的应用前景。

机器学习算法在医疗行业领域疾病诊断和健康风险监测中的应用发展迅速，基于机器学习算法开发的预测模型和智能健康管理系统，可以在线对疾病复发筛查数据的远程监测、实时更新和智能健康预警[29]。机器学习算法在获取数据体征与分析复杂数据方面具有显著优势，从早期经典的逻辑回归统计、决策树分析等传统方法，到新兴的深度神经网络学习算法，机器学习已被广泛应用于各个领域[30]。SVM 是机器学习中的常用算法，其强大的学习功能发挥着重要优势。饶飘雪等[31]采用 SVM、BPNN、Logistic 回归建立乳腺癌复发预测模型，结果显示，SVM 和 BPNN 模型具有更好的预测性，本研究结果与之相似。考虑到 DFUs 丰富的临床数据在复发风险预测中的重要意义，若能有效利用上述风险预测模型进行分析，提前控制高危因素，可避免或延缓疾病的不良进展。此外，不同模型具有各自的优势，与 SVM 模型相比，BPNN 模型存在局部极小值和收敛速度慢等不足，与 Logistic 回归模型相比，SVM 模型不直接依赖于数据分布，而对于多维、非线性医疗数据，SVM 在分类性能、泛化能力、建模计算量等方面具有明显优势[32]；Logistic 回归模型中的回归系数虽可解释各个影响因素的流行病学意义，但是却无法对存在严重共线性的变量作出明确解释[33-34]。

综上所述，基于 SVM 的 DFUs 复发风险预测模型效果最优，运用于 DFUs 临床评估将有助于提高易复发人群筛查率，可以提前给予复发风险预警，利于降低疾病负担且节约医疗成本，为 DFUs 防治提供一定的临床指导和帮助。除此之外，在临床科研方面，通过预测复发风险实现对患者的分组，从而控制组间的可比性。由于本研究仅关注了 DFUs 的短期预后结局，今后可扩充样本量，纳入不同地域的人群，开展多中心和大样本的队列研究，探索新的输入变量，优化模型准确度，进一步验证预测模型的效能。采用多种机器学习算法，如决策树和随机森林等进行比较，联合预测以综合判断模型的预测价值，构建更权威的模型算法提升研究的准确性和推广性。

作者贡献：张娟提出研究思路，设计研究方案，研究的实施与可行性分析，包括复发风险预测模型建模思路，数据收集，文章的构思和论文撰写；张娟、李海芬，李小曼、姚苗、马惠珍负责筛选研究对象，数据采集和整理，构建预测模型统计学处理，结果的分析与解释，图表的设计和制作；马强负责论文的修订，负责文章的质量控制及审校，提供研究经费及材料支持，对文章整体负责，监督管理。

本文无利益冲突。

参考文献

[1] SAEEDI P，PETERSOHN I，SALPEA P，et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045：results from the International Diabetes Federation Diabetes

Atlas，9th edition［J］. Diabetes Res Clin Pract，2019，157：107843. DOI：10.1016/j.diabres.2019.107843.

［2］TENTOLOURIS N，EDMONDS M E，JUDE E B，et al. Editorial：understanding diabetic foot disease：current status and emerging treatment approaches［J］. Front Endocrinol（Lausanne），2021，12：753181. DOI：10.3389/fendo.2021.753181.

［3］ZHOU X Q，LI M，XIAO M F，et al. ERβ accelerates diabetic wound healing by ameliorating hyperglycemia-induced persistent oxidative stress［J］. Front Endocrinol（Lausanne），2019，10：499. DOI：10.3389/fendo.2019.00499.

［4］CAVANAGH P R，LIPSKY B A，BRADBURY A W，et al. Treatment for diabetic foot ulcers［J］. Lancet，2005，366（9498）：1725-1735. DOI：10.1016/s0140-6736(05)67699-4.

［5］MCGLOIN H，DEVANE D，MCINTOSH C D，et al. Psychological interventions for treating foot ulcers，and preventing their recurrence，in people with diabetes［J］. Cochrane Database Syst Rev，2021，2（2）：CD012835. DOI：10.1002/14651858. CD012835.pub2.

［6］YADAV C，SRIKANTIAH R M，MANJREKAR P，et al. Assessment of mineral pathophysiology in patients with diabetic foot ulcer［J］. Biol Trace Elem Res，2020，195（2）：366-372. DOI：10.1007/s12011-019-01868-3.

［7］ARMSTRONG D G，BOULTON A J M，BUS S A. Diabetic foot ulcers and their recurrence［J］. N Engl J Med，2017，376（24）：2367-2375. DOI：10.1056/NEJMra1615439.

［8］吕静，袁丽，李饶，等. 糖尿病足溃疡复发风险预测模型的构建［J］. 护理研究，2022，36（6）：993-998. DOI：10.12102/j.issn.1009-6493.2022.06.009.

［9］CRAWFORD F，CEZARD G，CHAPPELL F M，et al. The development and validation of a multivariable prognostic model to predict foot ulceration in diabetes using a systematic review and individual patient data meta-analyses［J］. Diabet Med，2018，35（11）：1480-1493. DOI：10.1111/dme.13797.

［10］AAN DE STEGGE W B，SCHUT M C，ABU-HANNA A，et al. Development of a prediction model for foot ulcer recurrence in people with diabetes using easy-to-obtain clinical variables［J］. BMJ Open Diabetes Res Care，2021，9（1）：e002257. DOI：10.1136/bmjdrc-2021-002257.

［11］《中国老年型糖尿病防治临床指南》编写组. 中国老年2型糖尿病防治临床指南（2022年版）［J］. 中国糖尿病杂志，2022，30（1）：10-11. DOI：10.3969/j.issn.1006-6187.2022.01.002.

［12］BUS S A，LAVERY L A，MONTEIRO-SOARES M，et al. Guidelines on the prevention of foot ulcers in persons with diabetes（IWGDF 2019 update）［J］. Diabetes Metab Res Rev，2020，36（Suppl 1）：e3269. DOI：10.1002/dmrr.3269.

［13］HINCHLIFFE R J，FORSYTHE R O，APELQVIST J，et al. Guidelines of the international writing group on the diabetic foot on diagnosis［J］. Diabetes/Metabolism Research and Reviews，2020，e3276. DOI：10.1002/dmrr.3276.

［14］李柏樟. 白癜风共病糖尿病风险预测模型的构建及验证［D］. 西安：空军军医大学，2022.

［15］CHIN Y F，HUANG T T. Development and validation of a diabetes foot self-care behavior scale［J］. J Nurs Res，2013，21（1）：19-25. DOI：10.1097/jnr.0b013e3182828e59.

［16］方蕾. 慢性病病人风险感知问卷的编制及其影响因素研究［D］. 西安：第四军医大学，2015.

［17］宋媛媛，陈兵，刘静静，等. BP神经网络模型与Logistic回归模型在精神分裂症患者复发影响因素分析中的比较［J］. 护理学报，2022，29（13）59-65. DOI：10.16460/j.issn1008-9969.2022.13.059.

［18］马倩倩，孙东旭，石金铭，等. 基于支持向量机与XGboost的成年人群肿瘤患病风险预测研究［J］. 中国全科医学，2020（12）：1486-1491. DOI：10.12114/j.issn.1007-9572.2020.00.066.

［19］PEIKES D，CHEN A，SCHORE J，et al. Effects of care coordination on hospitalization，quality of care，and health care expenditures among Medicare beneficiaries：15 randomized trials［J］. JAMA，2009，301（6）：603-618. DOI：10.1001/jama.2009.126.

［20］WANG L M，GAO P，ZHANG M，et al. Prevalence and ethnic pattern of diabetes and prediabetes in China in 2013［J］. JAMA，2017，317（24）：2515-2523. DOI：10.1001/jama.2017.7596.

［21］JIANG Y F，WANG X M，XIA L，et al. A cohort study of diabetic patients and diabetic foot ulceration patients in China［J］. and，2015，23（2）：222-230. DOI：10.1111/wrr.12263.

［22］HICKS C W，CANNER J K，MATHIOUDAKIS N，et al. Incidence and risk factors associated with ulcer recurrence among patients with diabetic foot ulcers treated in a multidisciplinary setting［J］. J Surg Res，2020，246：243-250. DOI：10.1016/j.jss.2019.09.025.

［23］FU X L，DING H，MIAO W W，et al. Global recurrence rates in diabetic foot ulcers：a systematic review and meta-analysis［J］. Diabetes Metab Res Rev，2019，35（6）：e3160. DOI：10.1002/dmrr.3160.

［24］中国微循环学会糖尿病与微循环专业委员会，中国研究型医院学会创面防治与损伤组织修复专业委员会. 糖尿病足基层筛查与防治专家共识［J］. 中国糖尿病杂志，2019，27（6）：401-407. DOI：10.3969/j.issn.1006? 6187.2019.06.001.

［25］夏磊，庄若，吴玲，等. 复发性糖尿病足溃疡相关危险因素的Meta分析［J］. 中华现代护理杂志，2022，28（9）：1143-1148. DOI：10.3760/cma.j.cn115682-20210802-03422.

［26］周罗晶，刘昕，高巨. 无痛结肠镜检查术患者术后疲劳综合征人工神经网络预测模型的建立［J］. 中华麻醉学杂志，2021，41（4）：397-400. DOI：10.3760/cma.j.cn131073.20201202.00403.

［27］孟萌，关玉珠，郭利敏，等. 心血管外科术后患者ICU后综合征风险预测模型的构建及验证［J］. 中华护理杂志，2022，57（12）：1486-1494. DOI：10.3761/j.issn.0254-1769.2022.12.012.

［28］王洁，杨永忠，郑子薇，等. 石油工人代谢综合征的三种风险预测模型比较［J］. 现代预防医学，2021，48（4）：613-618. DOI：10.7620/zgfybj.j.issn.1001-4411.2017.23.03.

［29］周丽娟，温贤秀，蒋蓉，等. 机器学习在护理领域中的应用研究进展［J］. 护士进修杂志，2022，37（15）1388-1392. DOI：10.16821/j.cnki.hsjx.2022.15.009.

［30］吴秋硕，陆宗庆，刘瑜，等. 2010至2020年机器学习方法在重症医学领域应用进展的可视化分析［J］. 中国急救医学，2021，41（11）：959-965. DOI：10.3969/j.issn.1002-

1949.2021.11.008.
[31] 饶飘雪，叶枫 . 基于 Logistic 回归、ANN、SVM 的乳腺癌复发影响因素研究［J］. 计算机系统应用，2016，25（7）：259-263. DOI：10.15888/j.cnki.csa.005181.
[32] ZHAN J B，WANG Q，CHENG J，et al. Current status of severe fever with thrombocytopenia syndrome in China［J］. Virol Sin，2017，32（1）：51-62. DOI：10.1007/s12250-016-3931-1.
[33] 汪奇，刘尚全 . 2 型糖尿病患者亚临床动脉粥样硬化的多层人工神经网络分类预测模型的构建［J］. 中国全科医学，2021，24（36）：4612-4617. DOI：10.12114/j.issn.1007-9572.2021.00.537.
[34] 覃伟，高敏，沈莹，等 . 基于机器学习算法的 2 型糖尿病患者 3 个月血糖预测［J］. 中华疾病控制杂志，2019，23（11）：1313-1317. DOI：10.16462/j.cnki.zhjbkz.2019.11.003.

（收稿日期：2023-03-22；修回日期：2023-05-23）
（本文编辑：康艳辉）